The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07)

A NOVEL MAXIMUM LIKELIHOOD ESTIMATOION OF SUPERIMPOSED EXPONENTIAL SIGNALS IN NOISE AND ULTRA-WIDEBAND APPLICATION

Hai Zhan[1,2] John Farserotu[1] Jean-Yves Le Boudec[2]
{hai.zhan, john.farserotu}@csem.ch {jean-yves.leboudec}@epfl.ch
[1]Wireless Communication Section, CSEM SA, CH-2007 Neuchatel, Switzerland
[2]LCA2,EPFL, Ecublens, CH-1015 Lausanne, Switzerland

ABSTRACT

We pose the estimation of the parameters of multiple superimposed exponential signals in additive Gaussian noise problem as a Maximum Likelihood (ML) estimation problem. The ML problem is very non linear and hard to solve. Some previous works focused on finding alternative estimation procedures, for example by denoising. In contrast, we tackle the ML estimation problem directly. First, we use the same transformation as the first step of Iterative Quadratic Maximum Likelihood (IQML) and transform the ML problem into another optimization problem that gets rid of the amplitude coefficients. Second, we solve the remaining optimization problem with a gradient descent approach ("pseudo-quadratic maximum likelihood"). We also use this algorithm for Ultra-Wideband channel estimation and estimate ranging in non-line of sight environment.

I. INTRODUCTION

The estimation of the parameters of multiple superimposed exponential signals in additive Gaussian noise is of interest in time series analysis and system identification and in antenna array processing.

Based on the ML formulation, we developed a new estimation algorithm that outperforms the existing ones. [1] formulated the ML criterion which a multidimensional cost function with many local minima including the global minimum, which is of greatest interest since it leads to the smallest estimation error that can be obtained. Using directly the obtained cost function would need adaptive techniques (with a big number of iterations). Our idea was to force the cost function to be quadratic in terms of the parameters to be estimated. The known techniques [1] (IQML) using this approach leads to good estimates for high values of the Signal to Noise Ratio (SNR) (about more than 50dB). At low values of the SNR, the proposed techniques diverge and the global ML cost function minimum is not found. We showed that this in-ability to find the global ML minimum is essentially due to the presence of noise in the ML cost function which biases the obtained estimates. We proposed an original estimation algorithm which leads the global ML minimum. The derivation of the algorithm is based on the proper formulation of the gradient of the ML cost function and its global convergence is in two iterations (a substantial reduction in process complexity compared to adaptive ML techniques) and its initialization is done implicitly (we do not need an explicit initialization using another estimation procedure). The obtained performance of the proposed estimation algorithm is quasi-optimal (since it resolves the noiseless ML optimization criteria) and its processing complexity is very low: only two iterations are requested to obtain a very accurate position estimate.

We also use this algorithm for Ultra-Wideband channel estimation and estimate ranging in non-line of sight environment.

Section II introduces our system model. In Section III we transform our maximum likelihood estimation problem into a pseudo-quadratic one. In Section IV we define our proposed pseudo-quadratic maximum likelihood algorithm. In Section V we use our algorithm for Ultra-wideband channel estimation and ranging. Section VI shows simulation results.

II. SYSTEM MODEL

Suppose that plane waves from L point sources from distinct directions impinge on a M element line array. The M signal samples at j^{th} instant in time are embedded in a vector $\vec{y}_j$ called snapshot vector [7].

$$\vec{y}_j = \left[y_j(0), y_j(1), \cdots, y_j(M-1)\right]^T \qquad (1)$$

A set of J measured data vectors $\vec{y}_j$, $j = 1, \cdots, J$,is available. The observed samples consists of signals from the L point sources, $\vec{w}_j = \left[w_j(0), w_j(1), \cdots, w_j(M-1)\right]^T$ and noise $\vec{N}_j = \left[n_j(0), n_j(1), \cdots, n_j(M-1)\right]^T$. That is

$$\vec{y}_j = \vec{w}_j + \vec{n}_j \qquad j = 1, \cdots, J \qquad (2)$$

The signal samples at the k^{th} element at the j^{th} snapshot is modelled as follows [6].

$$w_j(k) = \sum_{l=1}^{L} a_{lj} \exp\left(j\frac{2\pi d}{\chi}\left(k - \frac{M+1}{2}\right)\sin\theta_l + j\rho_{lj}\right) \qquad (3)$$

where L : number of point sources

d :spacing between elements

χ : wavelength of radiation

M : number of elements in the array: $M > 2L$

θ_l : Bearing of the l^{th} source

a_{lj} : Amplitude of the l^{th} source at the j^{th} snapshot

ρ_{lj} : Phase angle of the l^{th} source at the j^{th} snapshot.

In a simplified form the signal sample are written as

$$w_j(k) = \sum_{l=1}^{L} a_{lj}\lambda_l^k \qquad k = 0, \cdots, M-1 \qquad (4)$$

W here $\lambda_l = \exp(jf_l)$ and $f_l = (d/\chi)\sin\theta_l$ and assume it is distinct from the other signals ,i.e, $\lambda_{l_1} \neq \lambda_{l_2}, \forall l_1 \neq l_2$. a_{lj} is the complex amplitude of the l th signal in the j th snapshot, which may vary across j .

The M components of the j th vector y_j are given by

$$y_j(k) = \sum_{l=1}^{L} a_{lj} \lambda_l^k + n_j(k) \quad k = 0, \cdots, M-1 \tag{5}$$

where $n_j(k)$ are complex normal random variables uncorrelated across both k and j ,with uncorrelated real and imaginary components, each of variance $\sigma^2/2$.

We can write (5) in the following vector notation.

$$\vec{y}_j = V(\vec{\lambda})\vec{a}_j + \vec{N}_j \tag{6}$$

where

$$\vec{y}_j = [y_j(0), y_j(1), \cdots, y_j(M-1)]^T$$
$$\vec{\lambda} = [\lambda_1, \lambda_2, \cdots, \lambda_L]$$
$$v(\lambda_l) = [1, \lambda_l, \lambda_l^2, \cdots, \lambda_l^{M-1}]$$
$$V(\vec{\lambda}) = [v(\lambda_1), v(\lambda_2), \cdots, v(\lambda_L)]$$
$$\vec{a}_j = [a_{1j}, a_{2j}, \cdots, a_{Lj}]^T$$
$$\vec{N}_j = [n_j(0), n_j(1), \cdots, n_j(M-1)]^T$$

The superscript ^{T} refers to the transpose operator and ^{H} refers to the Hemitian-transpose operator.

The above formulation encompasses the data model for a variety of problems. The case $J = 1$ corresponds to the single experiment time series problem with uniform sampling, the components of the vector $\vec{y} = \vec{y}_1$ being the M time samples of the measured signal. The case $J > 1$ corresponds to a multiple experiment with a time series, or equivalently to a multiple experiment with a time series, or equivalently, to the data model for a linear uniform narrow-band array, with multiple plane waves (far-field sources) present. In the latter case, the components of the measurement vector represent the output from M individual elements.

Assume that the number of signals L is known, given the data $\{\vec{y}_j\}_{j=1}^J$,we need to estimate the signal parameter vector $\vec{\lambda}$ and the signal amplitude $\{\vec{a}_j\}_{j=1}^J$. Finally, we can compute bearing of the source $\vec{\theta} = [\theta_1, \theta_2, \cdots, \theta_L]^T$ with $\vec{\lambda}$.

III. MAXIMUM LIKELIHOOD ESTIMATION

When the observation noise is Gaussian, the Maximum Likelihood (ML) criterion is equivalent to a Least Squares (LS) one. The ML estimate of the signal parameters $\vec{\lambda}$ and amplitudes $\{\vec{a}_j\}_{j=1}^J$ is obtained by solving the nonlinear least squares problem

$$\min_{\vec{\lambda}, a_j} \sum_{j=1}^{J} \left\| \vec{y}_j - V(\vec{\lambda})\vec{a}_j \right\|^2 \tag{7}$$

where $\|\cdot\|$ represents the Euclidean norm.

Assume that we pick a $V(\vec{\lambda})$, the best $\vec{a}_j$ for that choice of $V(\vec{\lambda})$ is obtained by

$$\hat{\vec{a}}_j = V(\vec{\lambda})^+ \vec{y}_j \tag{8}$$

where $V(\vec{\lambda})^+ = [V^H(\vec{\lambda})V(\vec{\lambda})]^{-1} V^H(\vec{\lambda})$ is the pseudo inverse of $V(\vec{\lambda})$. By substituting (8) into (7), the problem (7) is reduced to the following formulations:

$$\min_{\vec{\lambda}} \left\{ tr\left[P_V^{\perp}(\vec{\lambda}) \hat{R}_y \right] \right\} \tag{9}$$

where $P_V^{\perp}(\vec{\lambda}) = I - V(\vec{\lambda})[V^H(\vec{\lambda})V(\vec{\lambda})]^{-1} V^H(\vec{\lambda})$, tr is the trace operator, and $\hat{R}_y = \frac{1}{J}\sum_{j=1}^{J} \vec{y}_j \vec{y}_j^H$ is the sample autocorrelation matrix of $\vec{y}_j$. Once the ML estimate $\hat{\vec{\lambda}}$ is determined by solving (9), $\hat{\vec{a}}_j$ are found by the (8).

As $V(\vec{\lambda})$ is a Vanderonde matrix, there exists a Toeplitz matrix, B , of dimension $M \times (M - L)$ such that

$$B^H V(\vec{\lambda}) = 0 \tag{10}$$

The matrix B is given by

$$B = \begin{bmatrix} b_0 & 0 & & \\ b_1 & b_0 & \ddots & \\ \vdots & b_1 & \ddots & 0 \\ b_L & \vdots & \ddots & b_0 \\ 0 & b_L & \ddots & b_1 \\ & \ddots & \ddots & \vdots \\ & & 0 & b_L \end{bmatrix} \tag{11}$$

and its elements are taken from the coefficients of polynomial

$$b(z) = b_0 z^L + b_1 z^{L-1} + \cdots + b_L \tag{12}$$

with roots equals to $\{\lambda_1, \lambda_2, \cdots, \lambda_L\}$.

$\vec{b} = [b_0, b_1, \cdots, b_L]$ has following constraints:

1) $\vec{b} \neq 0$ 2) $\|\vec{b}\| = 1$ 3) $\vec{b} : b_i = b_{L-i}^*, i = 0, \cdots, L$

From Eq.(10), we know

$$P_V^{\perp} = P_B \tag{13}$$

We can write Eq.(9) as follows:

$$\min_{\vec{b}} \left\{ tr\left[P_B \hat{R}_y \right] \right\} = \min_{\vec{b}} \left\{ tr\left[B(B^H B)^{-1} B^H \hat{R}_y \right] \right\} \tag{14}$$

Once $\hat{\vec{b}}$ is obtained by Eq.(14), $\vec{\lambda}$ is obtained by the roots of $b(z)$.

Since B^H is linear in $\vec{b}^*$ (where the superscript * refers to the complex conjugate operation), a matrix Y_j filled out with the elements of $\vec{y}_j$ can be found such that $B^H \vec{y}_j = Y_j \vec{b}^*$. Eq.(14) becomes

$$\min_{\vec{b}} \frac{1}{J} \sum_{j=1}^{J} \left\{ \vec{b}^H Y_j R^{-T} Y_j^* \vec{b} \right\} \quad \text{where } R = B^H B \quad (15)$$

Because J is a constant, Eq.(15) is equivalent to

$$\min_{\vec{b}} \sum_{j=1}^{J} \left\{ \vec{b}^H Y_j R^{-T} Y_j^* \vec{b} \right\} \quad (16)$$

In iterative quadratic maximum likelihood (IQML) algorithm [1], the search $\vec{b}$ is replaced by an iterative procedure. For a given IQML iteration, the vector $\vec{b}$ in R^{-T} is held fixed, and the Eq.(16) is then quadratic in the remaining involving $\vec{b}$ and can be minimized in closed-form. The resulting $\vec{b}$ is then used to fix R^{-T}, and the process is repeated until $\vec{b}$ converges. IQML algorithm leads to good estimates for high values of the Signal to Noise Ratio (SNR). At low values of the SNR, the IQML approach diverge and the global ML cost function minimum is not found. In the next part, we used pseudo-quadratic maximum likelihood (PQML) algorithm to find the global ML cost function minimum in the low SNR environments.

IV. Pseudo-Quadratic Maximum Likelihood

In [2][3], the Pseudo Quadratic Maximum Likelihood (PQML) approach was shown to outperform proposed ML based multichannel estimation techniques and to lead to the ML cost function minimum. Following a development similar to the one in [4], it can be shown that the gradient of the ML cost function can be written as $\Gamma(\vec{b})\vec{b}$,where $\Gamma(\vec{b})$ is (ideally) positive semi-definite. A stationary point of the ML cost function satisfies $\Gamma(\vec{b})\vec{b} = 0\vec{b}$. This nonlinear eigenvalue problem is solved using the following iterative algorithm.

1. Choose a starting point $\vec{b}_0$
2. For each $\vec{b}_k$,construct $\Gamma(\vec{b}_k)$
3. Choose $\vec{b}_{k+1}$ to be the eigenvector corresponding to the smallest absolute eigenvalue of $\Gamma(\vec{b}_k)$
4. Repeat until convergence.

For our problem, $\Gamma(\vec{b})$ can be got from Eq.(16) as follows:

$$\Gamma(\vec{b}) = \sum_{j=1}^{J} Y_j^T R^{-T} Y_j^* - \sum_{j=1}^{J} \Psi_j^H \Psi_j \quad (17)$$

where matrix Ψ_j is defined as $T_j = R^{-1} B^H \vec{y}_j$ and $BT_j = \Psi_j \vec{b}$.

For finite M , the matrix $\Gamma(\vec{b})$ is indefinite and applying directly the PQML strategy will not work except for high SNR. We introduce an arbitrary γ in the minimization criterion, which becomes:

$$\min_{\vec{b},\gamma} \left\{ \vec{b}^H \left(\sum_{j=1}^{J} Y_j^T R^{-T} Y_j^* - \gamma \cdot \sum_{j=1}^{J} \Psi_j^H \Psi_j \right) \vec{b} \right\} \quad (18)$$

with semi-definite positive constraint on the central matrix.Hence, $\vec{b}$ can be found as the minimal generalized eigenvector of $\sum_{j=1}^{J} Y_j^T R^{-T} Y_j^*$ and $\sum_{j=1}^{J} \Psi_j^H \Psi_j$. γ is the corresponding minimum generalized eigenvalue. Asymptotically, there is a global convergence for $\vec{b}$, which converges to the ML minimum. Note also that the stationary points of PQML are the same as those of ML criterion; this is why PQML has the same performance as the ML criterion and hence gives the global ML minimum.

V. Ultra-Wideband Channel Estimation And Ranging

We will use the PQML algorithm for ultra-wideband channel estimation and ranging in this section. Consider a multipath channel with finite delay spread and an impulse response given by

$$h(t) = \sum_{l=1}^{L} a_l \delta(t - \tau_l) \quad (19)$$

where a_l and τ_l denote the amplitude and the time delays of the L channel paths, respectively. Define $\vec{\tau} = [\tau_1, \tau_2, \cdots, \tau_L]^T$ and $\vec{a} = [a_1, a_2, \cdots, a_L]^T$.

Assume that a sequence of Ultra-Wideband (UWB) pulses with waveform, $w(t)$ of duration T_p, is received through this channel. The received signal can be expressed as

$$r(t) = \sum_{l=1}^{L} a_l w(t - \tau_l) + n(t) \quad (20)$$

where $n(t)$ is an additive independent white Gaussian noise with zero mean and variance equal to σ^2. We also assume the transmitter and the receiver are synchronized.

We are interested in the joint estimation of the $\vec{\tau}$ and $\vec{a}$. Assume we receive Q samples, concatenated in the vector $\vec{r}$ (Q satisfies $T_o = MT_s$ and T_s denotes the sampling rate) through the transmission of G samples forming the UWB pulse ($T_p = GT_s$ and $G < Q$).

Assuming that the signal has finite energy, taking the Fourier transform of the received signal and assuming that M samples of the Fourier transform are available at the set of frequencies $\{f_1, f_2, \cdots, f_M\}$,we obtain

$$\vec{r}_s = SV(\vec{\lambda})\vec{a} + \vec{N}_s \quad (21)$$

where $\vec{r}_s$ ($\vec{r}_s = [r_s(f_1), \cdots, r_s(f_M)]^T$) is the Fourier transform of the received signal. S is an $M \times M$ full rank diagonal matrix and its elements corresponds to the Fourier transform of the UWB pulse, $\vec{W}(f)$ taken at the set of frequencies $\{f_1, f_2, \cdots, f_M\}$ such that , $f_{k+1} = \dfrac{k}{M}$ and $k = 0,1,\cdots, M-1$. S is defined as

$$S = \begin{bmatrix} W(f_1) & & \\ & \ddots & \\ & & W(f_M) \end{bmatrix} \quad (22)$$

$V(\vec{\lambda})$ is an $M \times L$ matrix which is defined as follows:

$$\lambda_l = e^{-j\frac{2\pi}{M}\tau_l}$$
$$v(\lambda_l) = [1, \lambda_l, \lambda_l^2, \cdots, \lambda_l^{M-1}]$$
$$V(\vec{\lambda}) = [v(\lambda_1), v(\lambda_2), \cdots, v(\lambda_L)]$$

The elements of $\vec{N}_s$ will be additive, independent Gaussian variables.

Multiple S^{-1} on the both sides of Eq.(21),we can get:

$$\vec{y} = V(\vec{\lambda})\vec{a} + \vec{N} \quad (23)$$

where $\vec{y} = S^{-1}\vec{r}_s$ and $\vec{N} = S^{-1}\vec{N}_s$. $\vec{N}$ is still additive ,independent Gaussian noise.

The problem of Eq.(23) corresponds to the case $J = 1$ in Eq.(6).

We can estimate $\vec{\lambda}$ and $\vec{a}$ with PQML algorithm first. And use $\lambda_l = e^{-j\frac{2\pi}{M}\tau_l}$ to get $\vec{\tau}$.

In case of ranging for UWB, the parameter we are interested in for the purpose of ranging is τ_1, which is the smallest element in the Time of Arrival (TOA) vector estimate, $\hat{\vec{\tau}}$, that is, the path that arrives first among all the detected paths.

VI. SIMULATIONS

We only simulate ultra-wideband channel estimation and ranging with PQML algorithm in this section. We consider a received UWB signal composed of Impulse Radio (IR) pulses. Each pulse has duration 1ns in a total burst length of 100ns, which means 1 percent duty cycle. On the IR UWB pulse, we apply an FFT of length N=128. The sampling frequency is 2GHz. The used channel is a five-ray propagation channel with path amplitudes equal to $\vec{a} = [0.6 \quad 0.8 \quad 0.6 \quad 0.7 \quad 0.5]$ and delays $\vec{\tau} = [10ns \quad 20.5ns \quad 31ns \quad 43.5ns \quad 50ns]$. All the presented results are average over 100 Monte-Carlo trials and for a given trial the Normalized Minimum Square Error (NMSE) for a vector of parameters $\vec{u}$ is defined by $\left\|\vec{u} - \hat{\vec{u}}\right\|^2 / \left\|\vec{u}\right\|^2$, and $\hat{\vec{u}}$ refers to the vector of parameters estimate. The initialization of the PQML algorithm is done by putting the matrix equal R to identity.

We consider now a Gaussian pulse generated with a centre frequency $f_c = 5.12GHz$, according to the following shape: $s(t) = 4\pi\sqrt{2.7}\, f_c e^{-2(\pi f_c t)^2}$ (the duration and duty cycles are as previously described.

Figure 1 illustrates the obtained NMSE of delay estimation versus the number of iterations at different values of SNR. NMSE of Channel Amplitude Estimation versus the number of iterations at different values of SNR is showed in Figure 2. From Figure 1 and Figure 2, we can note that there is a significant performance improvement when using the PQML estimation algorithm. Figure 3 illustrates the obtained performance in terms of standard deviation of the ranging error and compares it to its Cramer-Rao Lower Bound (CRLB). We can conclude that using the PQML algorithm for the estimation of the UWB TOA and then translating this to a ranging estimation leads to good results for the UWB ranging. From the above simulation results, the PQML algorithm is better than Subspace algorithm. The reason is the PQML algorithm is based on Maximum Likelihood Estimation estimator which is optimal estimator under Gaussian noise environment.

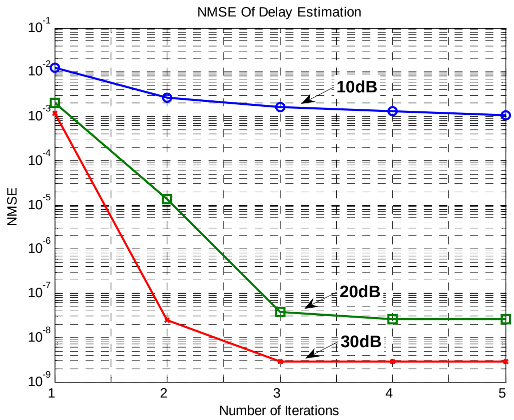

Figure 1: NMSE of Delay Estimation

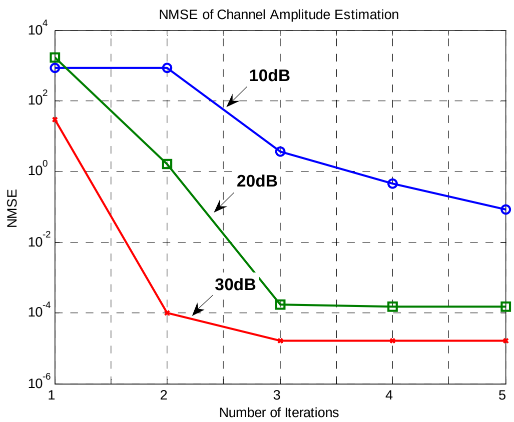

Figure 2: NMSE of Channel Amplitude Estimation

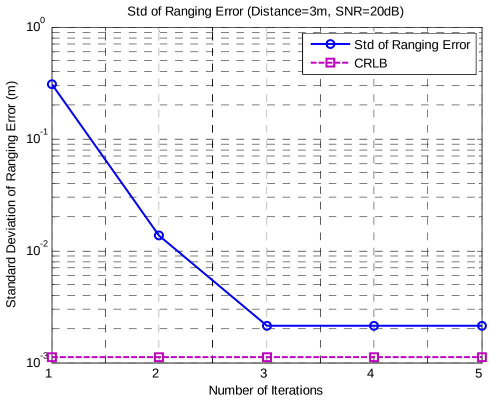

Figure 3: Standard Deviation of Ranging Error

VII. CONCLUSIONS

A high resolution iterative algorithm, PQML, has been proposed for the solution of the constrained optimization problem associated with the MLE requiring only the solution of linear equations. The algorithm can be used in Ultra-wideband channel estimation and ranging. We will also consider the cases of non Gaussian, impulsive interference noise in the future work.

REFERENCES

[1] Y.Bresler, A.Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise" *IEEE Transactions on Acoustics,Speech and Signal Processing,*, vol. ASSP-34, no. 5, October 1986.

[2] J. Ayadi, H. Zhan, J. Farserotu. "Maximum Likelihood Time of Arrival Esitmation for Ultrawideband Signals" *ISSPA2007*, Sharjah, United Arab Emirates, 12-15 February,2007

[3] J. Ayadi, E. de Carvalho,D.T.M. Slock. "Blind and Semiblind Maximum Likelihood Methods for FIR Multichannel Identification". *In Proc. ICASSP 98 Conf.*, USA, May 1998.

[4] G.Harikumar and Yoram Bresler, "Analysis and comparative evaluation of techniques for multichannel blind deconvolution". *SSAP '96*.

[5] M.R.Osborne and G.K.Smyth, "*A modified prony algorithm for fitting function defined by difference equations*", SIAM

[6] S.Haykin, "*Array Signal Processing*", Prentice-Hall,1985

[7] R.Kumaresan, A.K.Shaw, "*High Resolution Bearing Estimation without Eigen Decomposition*", ICASSP'85